Mirjalili *et al. Sustainable Energy Research (2023) 10:8*
https://doi.org/10.1186/s40807-023-00078-9

Sustainable Energy Research

ORIGINAL RESEARCH ARTICLE

Open Access

A comparative study of machine learning and deep learning methods for energy balance prediction in a hybrid building-renewable energy system

Mohammad Amin Mirjalili[1], Alireza Aslani[1*], Rahim Zahedi[1] and Mohammad Soleimani[1]

Abstract

Globally, the construction industry is experiencing an increase in energy demand, which has significant environmental and economic repercussions. To address these issues, it is now possible for buildings, vehicles, and renewable energy sources to collaborate and function as an advanced, integrated, and environmentally favorable system that meets the high energy demands of contemporary buildings. To attain maximum efficiency, however, it is necessary to create reliable energy demand forecasting models. In this research, by introducing the energy model of a neighbourhood with buildings with solar panels and electric vehicles, the final balance of energy production and consumption for each building and the whole neighbourhood as a micro grid is predicted. DesignBuilder is used to model neighbourhood buildings, and K-Nearest neighbor (KNN), Regression Support Vector (SVR), Adaptive Boosting (AdaBoost), and Deep neural networks (DNN) algorithms in machine learning are used to predict the final energy balance. a comparative analysis of the performance of the KNN, SVR, AdaBoost, and DNN algorithms was conducted to determine which algorithm is the most effective in predicting energy balance. Finally, the Root Mean Square Error (RMSE) has been used to validate the prediction models. The results show that the KNN, SVR, AdaBoost, and DNN algorithms had RMSE values of 0.56, 0.92, 0.95, and 0.53, respectively. Among these algorithms, the DNN and KNN algorithms had more accurate results than the other used algorithms and as a result of this research, An accurate forecast of neighbourhood energy balance was made. This study takes a novel approach by developing a model that takes into account an integrated system of houses, solar cells, and electric consumption for each building in a neighborhood, which can help to optimize energy consumption and reduce environmental impact.

Highlights

- Predicting the neighborhood energy balance considering PV and EV systems
- Analyzing the aerial and satellite maps of the neighborhood and extracting buildings types
- Selecting the best algorithm to predict the neighborhood energy balance from machine learning

Keywords Renewable energy, Electric vehicle, Forecasting, Final energy balance, Machine learning

*Correspondence:
Alireza Aslani
alireza.aslani@ut.ac.ir
Full list of author information is available at the end of the article

Introduction

The need to reduce global energy end-use and greenhouse gas emissions worldwide has boosted the implementation of energy-efficient opportunities (Env & IEA, 2017). According to the World Watch Institute, as the largest energy consumer, buildings account for 40% of annual global energy consumption and 36% of total carbon emissions, especially in urban areas (Jain et al., 2014; Wang et al., 2019). Growth rates of building energy consumption in countries (OECD[1]) and non-OECD countries for 2012 and 2040 are 1.5% and 1.2% per year, respectively (Conti et al., 2016). According to Canadian Natural Resources, the residential sector accounts for 13% of Canada's final consumption (Neves et al., 2017).

Renewable resources play an essential role in the decarbonization of the world community. To this end, it is well known that a great deal of effort must be made to redevelop cities through the reconstruction of the building and transportation sectors, which are responsible for most of the total energy consumption and CO_2 emissions (Lirola et al., 2017; O'Dwyer et al., 2019). In this regard, electric vehicles are a suitable solution to achieve zero carbon emissions. Also, the use of renewable energy sources along with buildings can effectively move towards buildings with high energy efficiency (Savvides et al., 2019). For this purpose, solar panels are recognized and used as the most popular renewable source. At the same time, the growth in the use of solar panels provides the opportunity to electrify part of transportation, and Electric vehicles can be charged using electrical energy from solar panels (Coffman et al., 2017). Electric Vehicle (EV) and Photovoltaic (PV) combined building design is expected to become more common to reduce overall electricity demand, which can lead to reducing fossil fuel consumption and greenhouse gas emissions in our cities. (Sawhney & Kahn, 2012) Very few papers have analyzed the connection of EV and PV in buildings (Buonomano et al., 2019; Quddus et al., 2018).

Energy management and control of PV and EV integrated systems refers to the process of optimizing the use of renewable energy produced by photovoltaic systems and integrating the charging and discharging of electric vehicles Energy management systems (EMS) can be used to monitor and control the energy flow in a PV and EV system that is integrated. An EMS can be used to optimize the utilization of renewable energy, manage energy storage, and regulate the charging and discharging of electric vehicles. The EMS can also be used to provide real-time information on energy consumption and system performance, enabling ongoing system optimization. Integration of PV and EV systems into the power grid effectively requires careful planning and coordination with local utilities. Grid integration entails ensuring that the system complies with all applicable local regulations and standards and is capable of importing and exporting energy to and from the grid as necessary. (Mokhtara et al., 2020, 2021).

Predicting the energy balance of houses integrated with PV and EV systems is an essential aspect of the development of sustainable energy solutions. While significant research has been conducted in this field, there are still some gaps that need to be filled. Although the integration of EVs with PV systems in residential buildings is gaining popularity, little is known about the interaction between EVs and PV systems. To investigate the impact of EV charging on the performance of PV systems, additional research is required. Numerous current models for predicting the energy balance of homes with PV and EV systems rely on simplified assumptions and models. There is a need for more data-driven models that can accurately predict and capture the complexity of these systems. Although integrating PV and EV systems with the grid is essential for the widespread adoption of these technologies, the impact of these systems on the grid is poorly understood. More research is required to investigate the impact of these systems on the grid and to develop management strategies for their interaction with the grid.

Buildings have become more intelligent from a smart grid perspective by combining advanced information and communication technologies, electric vehicles, decentralized storage systems, and energy production and management systems. Recently, research into forecasting energy consumption in buildings has become increasingly significant. Accurate and reliable energy demand forecasts enable large companies to plan resources and balance supply and demand, thus ensuring the electricity grid's stability and security and services' reliability (Deb et al., 2017). Therefore, energy consumption forecasting models to improve the energy efficiency of buildings for a sustainable economy have become an integral part of the building energy management system (BEMS[2]) (Xiao et al., 2018). Energy consumption forecasting models are generally divided into three categories:

- *Engineering methods* uses physical and thermodynamic laws and requires complex building and environmental parameters that are often time-consuming, such as Energyplus (Bui et al., 2020).
- *Statistical methods* Relate energy consumption to related factors such as climate data, housing. It lacks

[1] Organization for Economic Co-operation and Development.

[2] Building Energy Management System.

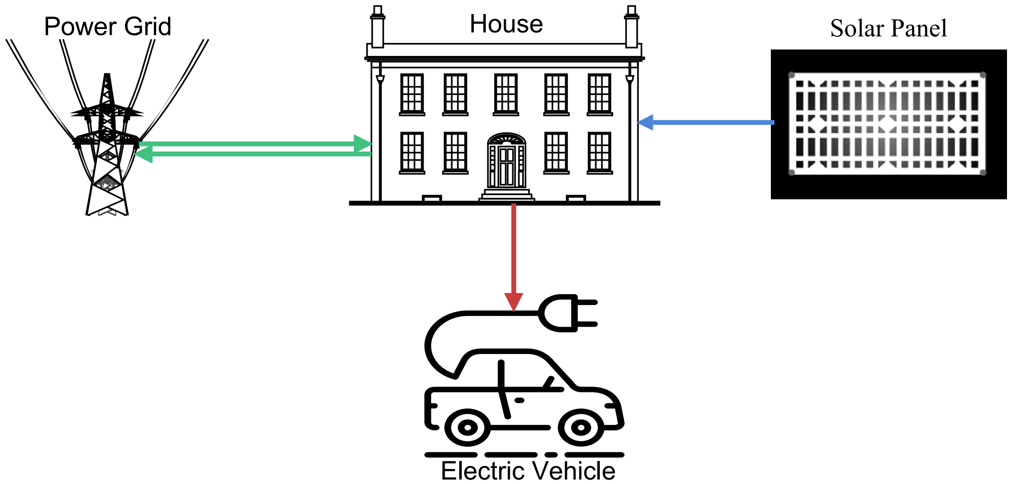


Fig. 1 Schematic of the selected system of buildings with EV and PV

precision and flexibility. Example—time series and linear regression (Deb et al., 2017)

- *Artificial intelligence methods* learns consumption patterns from historical energy consumption data, for example, discovering a nonlinear relationship between input (historical data) and output (target consumption) (Liu et al., 2020a), for example, the SVR algorithm. Meanwhile, artificial intelligence approaches have become an "active research hub" due to their efficiency and flexibility compared to engineering and statistical methods. This study has tried to predict the final energy balance model for an urban area in Edmonton. The buildings in this area, as in Fig. 1, together with PV and EV, form a single system.

This research's goal is to predict the neighbourhood energy balance considering PV and EV systems. First, some houses were selected as a case study in Edmonton, Canada, and by analyzing the aerial and satellite maps of this neighbourhood, the buildings types were extracted. The renewables.ninja algorithm predicts the generation of electricity from solar cells. The simulation output was given to the input of machine learning algorithms, and the best algorithm was selected to predict the neighbourhood energy balance.

Literature review

Electric vehicle

Given the growing concern about the effects of climate change, such as the loss of sea ice and rising sea levels, to severe events such as hurricanes, droughts, or extreme heatwaves, it is difficult to deny the dimensions of what we are struggling with here, and if we increase the temperature by two degrees Celsius on Earth, we will have serious consequences (Moss et al., 2010) .To minimize these consequences, scientists are investigating the leading causes of climate change. They found that greenhouse gases (GHG)[3] such as carbon dioxide, methane, nitrogen oxides, and aerosols change the atmosphere and affect the planet more. Since cars make up 72% of CO_2 emissions in the transportation sector (followed by aircraft with 10%), the electric car market is growing and seems to be an excellent solution to combat climate change (Gonçalves, 2018). Electric refueling gives advantages not found in ordinary cars with internal combustion engines. Electric vehicles use energy much more efficiently than internal combustion vehicles. This efficiency can significantly reduce the energy required in the transportation system and help achieve an independent future of fossil fuels. The most important advantage of electric cars is their contribution to improving air quality in cities and towns. Pure electric cars without exhaust pipes do not emit carbon dioxide while driving. This significantly reduces air pollution. While reducing CO_2 emissions can help reduce global warming and the many related adverse effects (sea-level rise, drought, and significant weather events), electric vehicles can also positively affect indoor air quality.

Different types of electric vehicles (Hampshire et al., 2018):

[3] Greenhouse gases.

- Battery-operated electric vehicles (BEV)[4] are powered solely by an electric motor and use electricity stored in an internal battery.
- Hybrid electric vehicles ((PHEV)[5] powered by an electric motor and an internal combustion engine run together or separately.
- Wide-range electric vehicles (REEV)[6] have a serial hybrid configuration in which their internal combustion engine has no direct connection to the wheels. Instead, the combustion engine acts as a generator when the electric motor is low, or the battery is recharged. The battery can also be charged from the mains.
- Hybrid electric vehicles (HEV)[7] combine an internal combustion engine and an electric motor to assist a conventional engine, for example, when accelerating a car.

In this research, the second type of EV is considered as a sample vehicle.

Zhang and his colleagues used deep learning algorithms to develop a traffic flow and electric intrusion model and evolved a model for charging an electric car (Zhang et al., 2020). Ramadhani and his colleagues have also studied a probabilistic load distribution model considering photovoltaic systems and electric vehicles, considering the uncertainty in distribution systems (Ramadhani et al., 2020). Guzel and His colleagues provide a way to develop unique charge models for PEV[8] users with core density estimates (KDE) for intelligent charging strategies (Guzel & Göl, 2021). Slama proposes an exact Home Centralized Photovoltaic (HOCP) System that incorporates V2H technology, Solar Photovoltaic (SPV), and a Green Electric Vehicle. The proposed concept aims to reduce domestic energy demand by offering optimal appliance automation (Zafar & Ben Slama, 2022). Khan studies the viability and design of a BIPV (building-integrated photovoltaics)-powered EV charging system in a typical Malaysian residence that uses solar energy to meet residential and EV charging requirements. Three BIPV systems have been constructed, simulated, and evaluated for their performance parameters: grid integrated with no battery, grid integrated with 75% battery storage, and grid integrated with 100% battery storage (Khan et al., 2022). Gholinejad proposes smart charging for off-board EV chargers in home-energy-hub (HEH) applications using solar and battery storage dc sources. Smart charging and discharging of EVs allows vehicle-to-x and x-to-vehicle operations for household applications linked with renewable and storage aspects (Gholinejad et al., 2022). Hou proposes a comprehensive strategy to prioritize user preferences while scheduling various physical equipment. A specific charging and discharging approach for the energy storage system and EV considering their capital cost is recommended to incorporate them into the HEMS for improved flexibility, economic benefits, and battery life. The smart home's energy schedule can be developed from mixed integer linear programming (MILP) and the proposed model to ensure user comfort and cheap cost (Hou et al., 2019).

The importance of machine learning and forecasting in the energy sector

Long-term forecasting of electricity consumption is the basis of energy investment planning and plays a vital role for developing countries. Accurate power modeling is critical to avoid costly mistakes (Conti et al., 2016). In a study of two commercial building tenants using SVM, ANN, and KNN techniques, M Shapi and his colleagues concluded that SVM has a minor error and more accurate prediction (Neves et al., 2017). Nutkiewicz and his colleagues report integrating a 3D engineering model of campus with 22 buildings and integrating it with forecasting algorithms. The final model has high accuracy and excellent dynamics that can be used in the future to predict the energy consumption of different urban contexts (Nutkiewicz et al., 2018). Xiaodong Xu and colleagues have developed a model for predicting the use of multi structural energy by combining social network analysis (SNA) with artificial neural network (ANN) techniques, and by dividing buildings into land uses, it has been able to make appropriate predictions with an accuracy of over 90% (Xu et al., 2019) .In another study, YangLiu developed a support vector method (SVM) for predicting and detecting energy consumption of public buildings based on 11 input parameters, including historical energy consumption data, climatic factors, and time cycle factors which can be used to manage building operations (Liu et al., 2020b (XJ. Luo and colleagues developed three machine learning multi-objective forecasting frameworks for simultaneously predicting multiple energy loads. Three machine learning techniques include artificial neural network, vector regression support, and short-term memory neural network. The ANN network-based prediction model has the lowest mean absolute error percentage, while an SVM-based one costs the shortest calculation time (Luo et al., 2020). Kim suggests a CNN-LSTM neural network to forecast dwelling energy use using spatial and temporal information.

[4] Battery electric vehicles.

[5] Plug-in hybrid electric vehicles.

[6] Range extended electric vehicles.

[7] Hybrid electric vehicles.

[8] Plug-in electric vehicles.

CNN-LSTM neural networks can extract complex energy usage information. The convolutional neural network (CNN) layer can extract features between energy-consuming variables, and the long short-term memory (LSTM) layer can represent temporal information of irregular trends in time series components (Kim & Cho, 2019).

Comprehensive information about Canada

It is difficult to track the quantity of off-grid capacity, which is deemed small in comparison to grid-connected capacity. In Canada, however, you can find off-grid solar PV applications (with or without battery storage) or hybrid systems that include a small wind turbine or diesel generator. These systems are often located in remote northern settlements, but as costs decrease and system installers and the public become more aware of the prospects, they are increasingly being built in less distant places. At the end of 2019, the national cumulative installed PV capacity was 3.33 GW_{DC}. This reflects an about 7.5% increase over the previous year. The growth in installed PV capacity in 2019 was 232 MW_{DC} (composed of 115.3 MW_{DC} of transmission-connected and 116.6 MW_{DC} distribution-connected capacity). Off-grid PV is not tracked and is presumed to be small in comparison to the total for grid-connected PV (Association, 2019). By 2030, Canada aims to reduce greenhouse gas emissions 30% below 2005 levels. To keep the world average temperature below 1.5 °C, Canada supports the 2015 Paris Agreement and the global transition to a low-carbon economy. Regarding PV policy, the definition of support measures is generally left to the provinces and territories. As previously stated, PV will be eligible for a number of national support programs announced by the Federal Government in 2017, including the $500 million Low Carbon Economy Challenge Fund, the $220 million Clean Energy for Rural and Remote Communities program, and the $100 million Smart Grid Program. (Association, 2019).

Detached houses, semi-detached houses, townhomes, and condominiums are the most common types of housing in Canada. 52.6% were single-detached houses, 5% were semi-detached houses, 6.5% were row-houses, 5.5% were apartments in a duplex, 18.3% were apartments in a building with fewer than 5 stories, 10.7% were apartments in a building with 5 or more stories, 0.2% were other single-attached houses, and 1.3% were mobile homes. Detached houses are one of the most common types of homes in Canada. A detached house is essentially a house that does not share a wall with another house and is surrounded on all sides by open space. In recent years, 2-Storey Houses have become the most prevalent type of detached home in Canada. Living and working areas are typically located on the lower level, while sleeping areas are typically located on the upper level. A bungalow is a type of detached house with all rooms on a single level. On occasion, a bungalow may have a study room or a bedroom on the attic level, which is accessible via stairs and built within the sloping roof. Split-level homes are homes in which the main floor is divided across multiple levels connected by a short flight of stairs rather than a full staircase. A 1.5-story home has two floors, but the second floor does not encompass the entire ground floor. Typically, the smaller space located upstairs is a bedroom ("All Types of Houses in Canada | WOWA.ca", 2023).

Methodology

Machine learning and explicitly programming solve computer science problems differently. Programmers write particular instructions or algorithms in explicitly programming. The program performs the programmer's commands in a predetermined order. The program can only handle specified input and output data and requires human involvement to update or modify. Machine learning programs learn from data without being programmed. Machine learning algorithms learn patterns and relationships from training data instead of describing each step. The trained algorithm can predict or decide on new data. More data helps the software and model improve. In essence, directly programming includes manually writing code to do a task, while machine learning involves training a program to learn from data and make predictions or choices. Both methods have pros and cons and suit different problems.

In order to implement this method, first, the studied case was examined and selected, and related information was collected. In the next step, information on the characteristics of buildings in the desired neighborhood and building standards in Canada was collected so that it could be used to simulate buildings in the DesignBuilder program. The output of building simulation and various building information was processed to be used as input for artificial intelligence tools to predict the balance of houses. The developed data served as input for the algorithms in the form of seven characteristics: hour, day, month, building type, solar panel surface area, building area for each building, and building heat load. The output and objective function of each model took each home's energy balance into account. In Fig. 2, the essential phases of the employed technique are outlined in detail.

After predicting each house's energy balance, the whole neighborhood's energy balance is calculated as a microgrid. At the end of this prediction, the types of methods used are checked for accuracy.

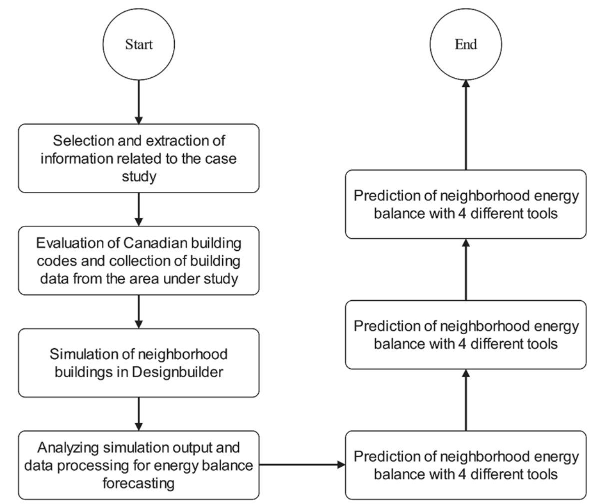


Fig. 2 Flowchart of detailing the essential stages of the employed method

Case study description

According to the purpose of the research, the use of building modeling tools and machine learning algorithms to measure the feasibility of using renewable energy to supply the energy needed for homes and electric vehicles in a neighbourhood has been investigated. A neighbourhood in Edmonton, Alberta, Canada, has been analyzed for this purpose. Figures 3 and 4.a show the 2D and 3D views of the case study, respectively.

The next step is the extent to which each house can absorb solar energy. Data on the amount of solar energy absorption potential and even the number of panels installed in each house in Edmonton is available on the MyHeat website (Solar-Myheat. 2023). The required information is available by referring to the solar maps section on this site. Figure 4b shows the solar map of the study area. By selecting the desired house, helpful information such as the maximum installable area of the panel according to the three-dimensional shape of the roof, the

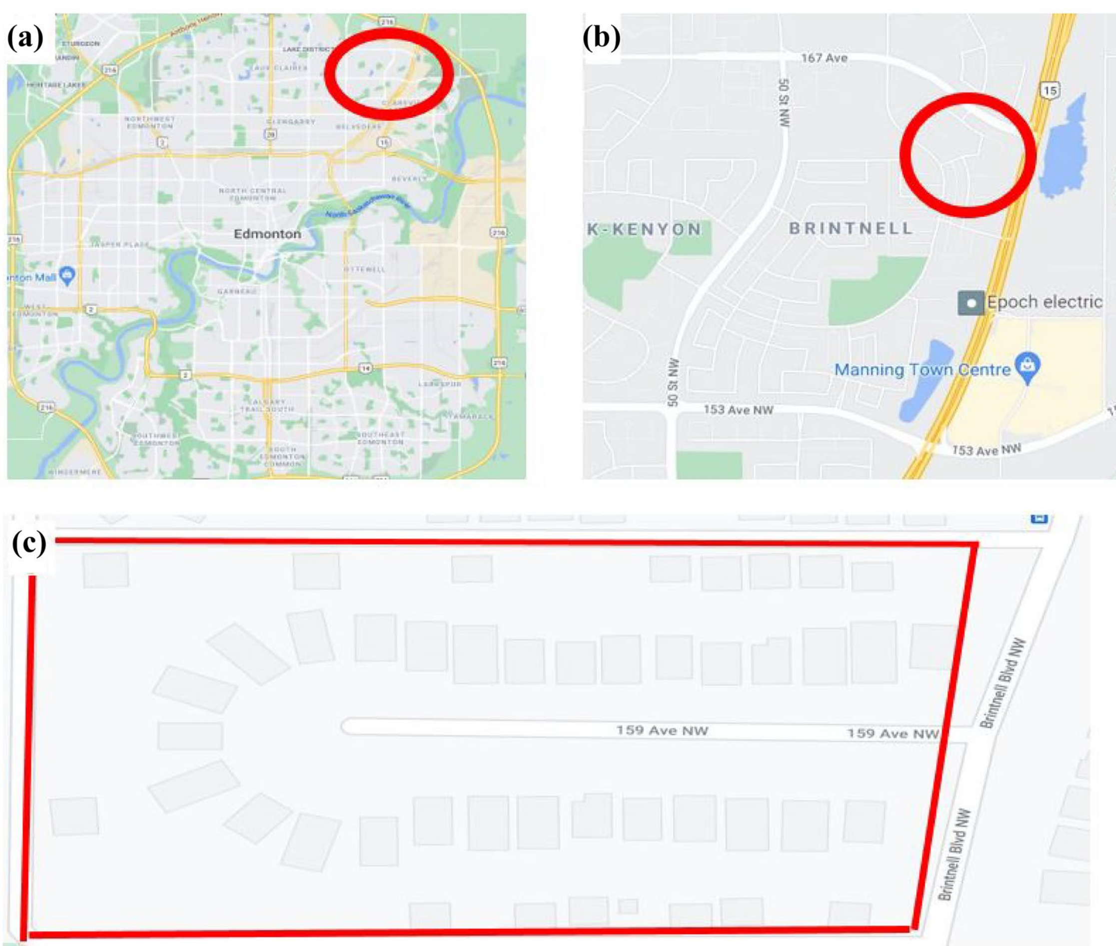


Fig. 3 **a** 2D view of the study area **b** Edmonton city map and selected neighbourhood in the northwest of the city **c** The study area in the BRINTNELL area

(a) (b)

Fig. 4 **a** 3D view of the case study **b** Solar potential of building roofs

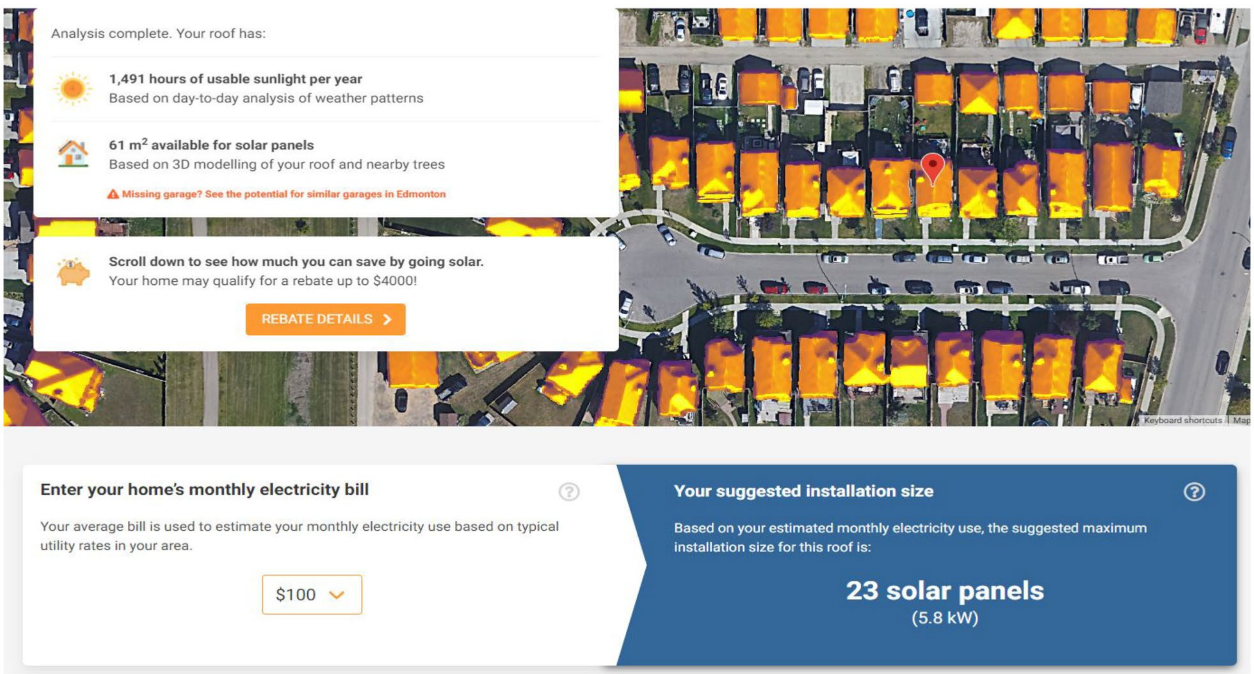


Fig. 5 MyHeat site output for a building located in the case study

number of required panels according to the site estimate, and the amount of net sunlight absorption during a year are displayed. Figure 5 shows a sample of the results.

Modelling in designBuilder

One of the most important parts of energy consumption in buildings is the heating sector. About 63% of energy consumption in Canadian buildings is spent on heating on average (City of Edmonton, 2017). Design Builder software is a powerful software with the Energy Plus engine that can calculate the amount of heating and cooling load. To obtain the desired output, we must first model the desired buildings in the software. Modeling in Design Builder software consists of three parts: (A) building geometry modeling. (B) entering the required information. (C) concluding.

In the first step of modeling, the drawings of the buildings were drawn in the Builder Design software. Since a specific map of the buildings in this neighbourhood is not available, the drawings on the Houseplans site have been used for modeling (HousePlans. 2018). The buildings were surveyed, and the types that most closely resembled the 3D views were selected and drawn. For example, Fig. 6 shows an example of these buildings and their map. In the next step, according to the exterior views of the buildings, four types of buildings on the Houseplans site were selected and simulated, and then the whole neighbourhood was modeled in the software. Figure 7 shows a view of the simulated neighbourhood in Builder Design. Then, the data related to Design Builder software, which is given as input, were collected and entered into the software. This data includes weather information of the area, information about the material and building materials and ventilation systems in each building. Edmonton's climate information is available by default in Design Builder, and it is sufficient to select it as a climatic reference for analysis at the beginning of the modeling. The selection of materials for the building, the type of layers, and their arrangement in different parts is very influential in the final result of the analysis. Since building information is not available in the selected area, the existing standards for construction in the area should be used to reduce the uncertainty in the results (Tian et al., 2018)

The Canadian National Building Standard (NBC[9]) is used to select layers. Section 9–36 of this standard contains tables that specify the minimum overall heat transfer coefficient for the various sections. Using the available resources for different parts of the conventional layering was selected according to the National Standard of Canada to comply with the items related to heat transfer. It should be noted that the city of Edmonton is known by the code 7a in the tables. Figure 7 shows an example of the materials selected for the exterior wall that meet the specified standards.

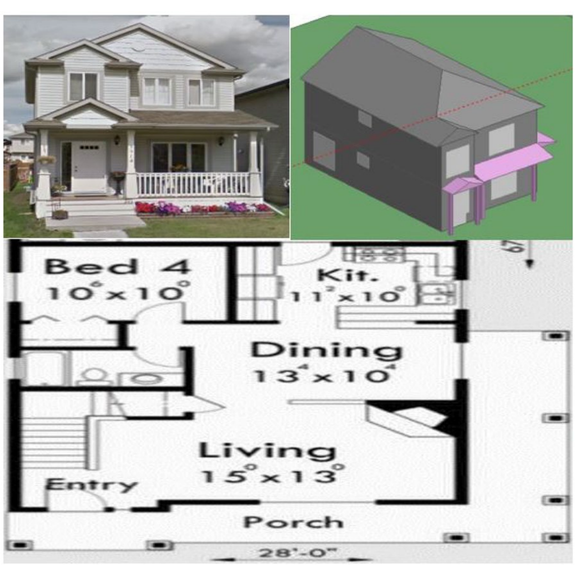


Fig. 6 Simulated buildings and maps of HousePlans

The software is ready to solve and present the results by entering the necessary information. The results will be displayed in the results section of the report. In this research, solar panels are used only to meet the demand for electrical appliances and electric vehicles, so the energy needs of the two energy consumers, space heating and water heating, are not met. Considering solar energy in the initial design is the best way to use solar energy for heating, and the use of panels is the best way to meet electricity demand.

Modelling building electrical load

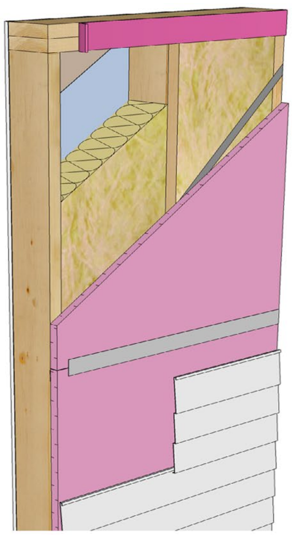
The assumptions made for calculating the electric charge of buildings in the neighbourhood are as follows. The number of people in each house was considered an average of 5 people. The minimum electrical appliances in each house are one TV, one refrigerator, one cooker, one washing machine, one microwave and mixer and tester, two laptops and four cell phones, three fans, and a hybrid electric car. According to the study area, the electrical charge output of each house is calculated on an hourly basis for 1 year using the simulation of DesignBuilder software. This calculation is without considering the amount of electric vehicle consumption. In the next step, the electric car's daily load is added to each house's household load, and finally, the total electricity consumption of each house is calculated in an hourly period for one year. For the electric vehicle intended for each household, the electric charge of the electric vehicle was extracted using the work method of Göhler and his colleagues, which is described in Fig. 8 (Göhler et al., 2019).

Figure 9 shows that the load peak is applied to the grid from 1 to 5 pm, as this time is approximately in the range of business hours for Canadian families.

By examining the best-selling solar panels in the Canadian market with the highest production efficiency, the SPR-A425-G-C model was selected, with the highest electricity generation efficiency of 22.8% among solar panels, by selecting the panel and considering the 22.8% efficiency rate and using the method developed by Pfenninger and his colleagues (Pfenninger & Staffell, 2016) for the potential of generating electricity using solar panels for each house and entering the exact geographical characteristics of the study area of each house. The basis of myheat site calculations is based on this article, which calculates the area that can be installed solar cells for each house by relying on the GIS system and choosing the geographical location of the houses. Hourly power per square meter was obtained as an annual data set.

[9] National Building Code of Canada.

Fig. 7 3D view of the analyzed neighbourhood and its simulated sample in Design Builder



	Thickness (mm)	RSI/mm	RSI-value
Exterior air film	-	-	0.03
Vinyl siding	-	-	0.11
Taped XPS Type 3	25	0.035	0.88
Metal let-in bracing	-	-	0.00
2x6 wood framing at 16" o.c. filled with R20 batt insulation*	-	-	2.36*
Polyethylene sheet	-	-	0.00
Gypsum board	12.7	0.0061	0.08
Interior air film	-	-	0.12
		Assembly effective RSI =	**3.58** (R-20.3)
		Zone 6 requirement (No HRV / HRV)	3.08 / 2.97
		Zone 7A requirement (No HRV / HRV)	3.08 / 2.97
		Zone 7B requirement (No HRV / HRV)	3.85 / 3.08

Fig. 8 The structure of external walls

Electricity generation calculations for each house are as follows:

$$P_{gen,i,h} = P_{gen,i,h,m} \times A_i \tag{1}$$

A_i $\mathrm{i_{th}}$ house area in the neighbourhood.

$P_{gen,i,h,m}$ $\mathrm{i_{th}}$ house Electricity generation capacity at the ith hour in 1 square meter.

$P_{gen,i,h}$ $\mathrm{i_{th}}$ house power generation capacity at $\mathrm{h_{th}}$.

The time period for reviewing the data is hourly, so the power and the amount of electricity consumed per hour are equal ($E_{i,h} = P_{i,h}$). Calculating the difference between production and demand shows what the condition of each house will be at each hour of the year. The hourly balance of total electricity for the house i at hour h is calculated as follows:

$$E_{demand,i,h} - E_{gen,i,h} = E_{Total,i,h} \tag{2}$$

$E_{demand,i,h}$ $\mathrm{i_{th}}$ house power consumption at $_{\mathrm{hth}}$ hour.

$E_{gen,i,h}$ $\mathrm{i_{th}}$ house power generation at $\mathrm{h_{th}}$ hour.

$E_{Total,i,h}$ $\mathrm{i_{th}}$ house electricity balance at hth hour.

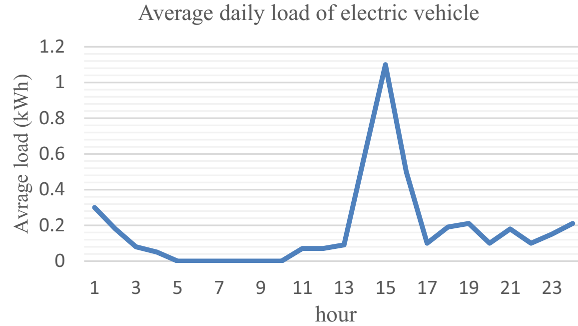


Fig. 9 Average daily electric vehicle load

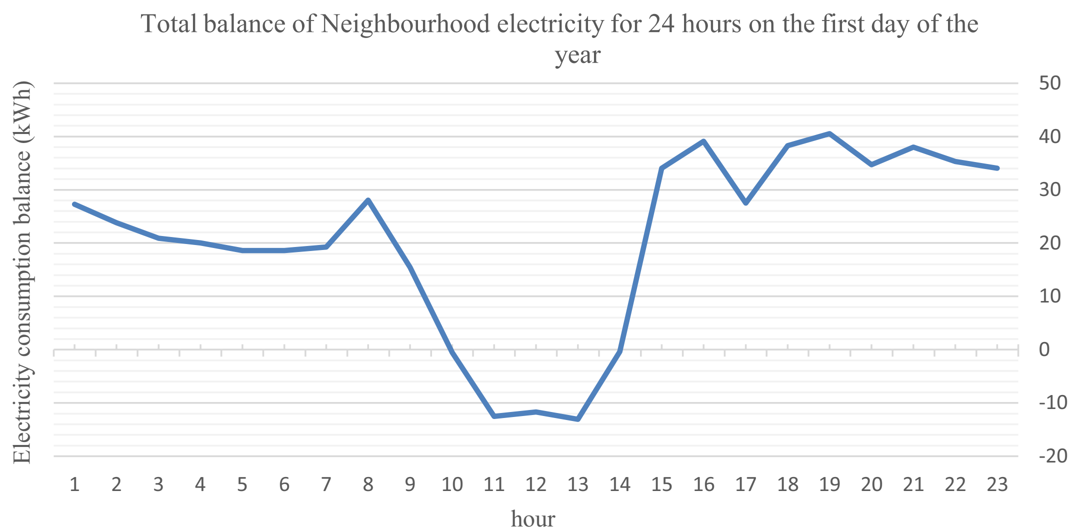


Fig. 10 Total electricity balance of the neighbourhood for 24 h on the 1 day of the year

If the electricity generation is more than the amount consumed, $E_{Total,i,h}$ will be negative, which can be used for other houses in the neighbourhood. The balance of total electricity for the neighbourhood is obtained from the following relation:

$$E_{Total,h} = \sum_{i=1}^{n} E_{Total,i,h} \qquad i \in 1, \ldots, n \tag{3}$$

n Number of houses in the neighbourhood

$E_{Total,h}$ Electricity balance of the whole neighbourhood at h hour.

If $E_{Total,h}$ at h hour is negative, the neighbourhood can export the generated electricity over its consumption to the general electricity grid.

Figure 10 shows the neighbourhood energy balance for 24 h from the 1 day of the year. As shown in the figure, the balance of energy consumption is negative in some hours, indicating that the excess electricity generated is delivered to the grid.

Machine learning

Machine learning and explicit programming are two distinct methods for solving computer science problems. In the explicitly programming methodology, programmers write specific instructions or algorithms to complete a task. The programmer specifies each phase of the program, which is then executed in a predetermined order. The resulting program is limited to specific input and output data and requires human intervention to be updated or adapted to new circumstances. In machine learning, the computer program is designed to automatically learn from data without being explicitly programmed. A machine learning algorithm is trained on a set of examples, known as training data, to identify patterns and relationships within the data. Once the algorithm has been trained, it can be used to predict or make decisions based on data it has never seen before. The program is capable of adapting and learning from new data, and the model's accuracy can increase over time.

In conclusion, explicitly programming requires the programmer to manually write code for a specific task, whereas machine learning involves training a program to learn from data and make predictions or decisions based on this learning. Both approaches have advantages and disadvantages, and are suited to various types of issues.

In the machine learning subject, the objective was defined so that by having the house characteristics, electricity consumption and production of the house, neighbourhood, and finally, the general condition of the electricity network can be calculated. According to the intended purpose, the existing problem is a regression problem, explained below. It is noteworthy that 20% of the data obtained are used for testing, and 80% are used for learning.

KNN algorithm

The first algorithm used for forecasting is the K-NN (K-Nearest neighbor) algorithm. This machine learning method is often used due to its simple criteria and predictability in a complex nonlinear pattern (Tian et al., 2018). This method predicts using similar items in the original data (Göhler et al., 2019). This method has been used to predict the energy balance of the building. Using optimization methods, the optimal value of K was 4.

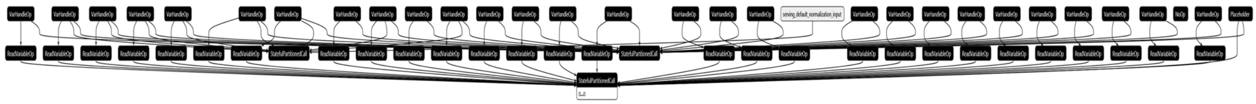


Fig. 11 Overview of DNN algorithm for the electric charge prediction problem

SVR algorithm

The support vector machine can be used as a regression method, retaining all the essential features that define the algorithm (maximum margin). Regression Support Vector (SVR) uses the same principles as SVM[10] for classification, with only a few minor differences. First, because the output is a real number, predicting the available information with infinite possibilities becomes complicated.

Adaboost algorithm

AdaBoost stands for Adaptive Boosting, a statistical classification algorithm formulated by Yoav Freund and Robert Schapire. It can be used with many other types of learning algorithms to improve performance. The output of other learning algorithms ("weak learner") is combined as a weighted sum representing the final output of the amplified classifier. AdaBoost is consistent in that subsequent poor learners change in favor of what has been done by previous classifiers. Some problems may be less prone to attachment than other learning algorithms. Individual learners can be weak, but as long as the performance of each is slightly better than random guessing, the final model can be converged to a strong learner. Each learning algorithm is more suited to some types of problem than others, and usually has different parameters and settings that must be adjusted before working optimally on a data set. AdaBoost with decision trees as (poor learners) is often known as the best out-of-the-box classification. When decision tree learning is used, the information gathered at each stage of the AdaBoost algorithm about each training instance's relative "hardness" is entered into the tree growth algorithm so that subsequent trees tend to be more challenging to classify.

ANN[11] algorithm

Artificial neural networks (ANN), commonly referred to as neural networks (NN[12]), are computational systems inspired by biological neural networks that make up the animal brain. ANNs are essentially massive parallel computational models that mimic the function of the human brain. An ANN consists of many simple processors interconnected by weighted connections. By analogy, processing nodes may be called "neurons". The output of each node depends only on information that is locally available in the node, whether stored internally or through weighted connections. Each unit receives inputs from many other nodes and sends its output to other nodes, which is not a very powerful processing element. It produces a scalar output with a single numeric value which is a simple nonlinear function of its inputs, and the system's power appears (Dongare et al., 2012). Deep neural networks (DNNs) are artificial neural networks (ANNs) with multiple concealed layers between the input and output layers. DNNs can model complex nonlinear relationships similarly to shallow ANNs. The primary function of a neural network is to receive a set of inputs, perform increasingly complex calculations on them, and produce output to solve real-world classification problems. We only consider feed forward neural networks. A deep network has an input, an output, and a flux of sequential data. According to the model developed for this problem, the number of input layer neurons is 128, and 64 hidden layers and one output layer are designed. Figure 11 shows the outline of the DNN algorithm implemented for the electric charge prediction problem.

Result and discussion

Design builder result

Design Builder software outputs both network and diagram. Its network output is an Excel file, which can be very efficient for analyzing data in machine learning algorithms. Also, these outputs can calculate the data related to energy consumption in different sectors on a monthly, daily, hourly, or even second level during a year.

The accuracy of software outputs is directly related to the accuracy of the data entered into the input nuances; Therefore, an essential part of modeling is the accuracy of the input data. In this research, the output is calculated for the amount of heating load and cooling load. Also, the amount of carbon dioxide production, which is one of the most critical environmental parameters, is received from the output. All of these outputs are for two consecutive years, calculated monthly basis each year. The outputs are also taken in the form of graphs and Excel files. Modeled buildings can be divided into three general categories based on geographical location:

- South–north buildings
- North–south buildings
- East–west buildings

[10] Support Vector Machine.

[11] Artificial Neural Network.

[12] Neural Network.

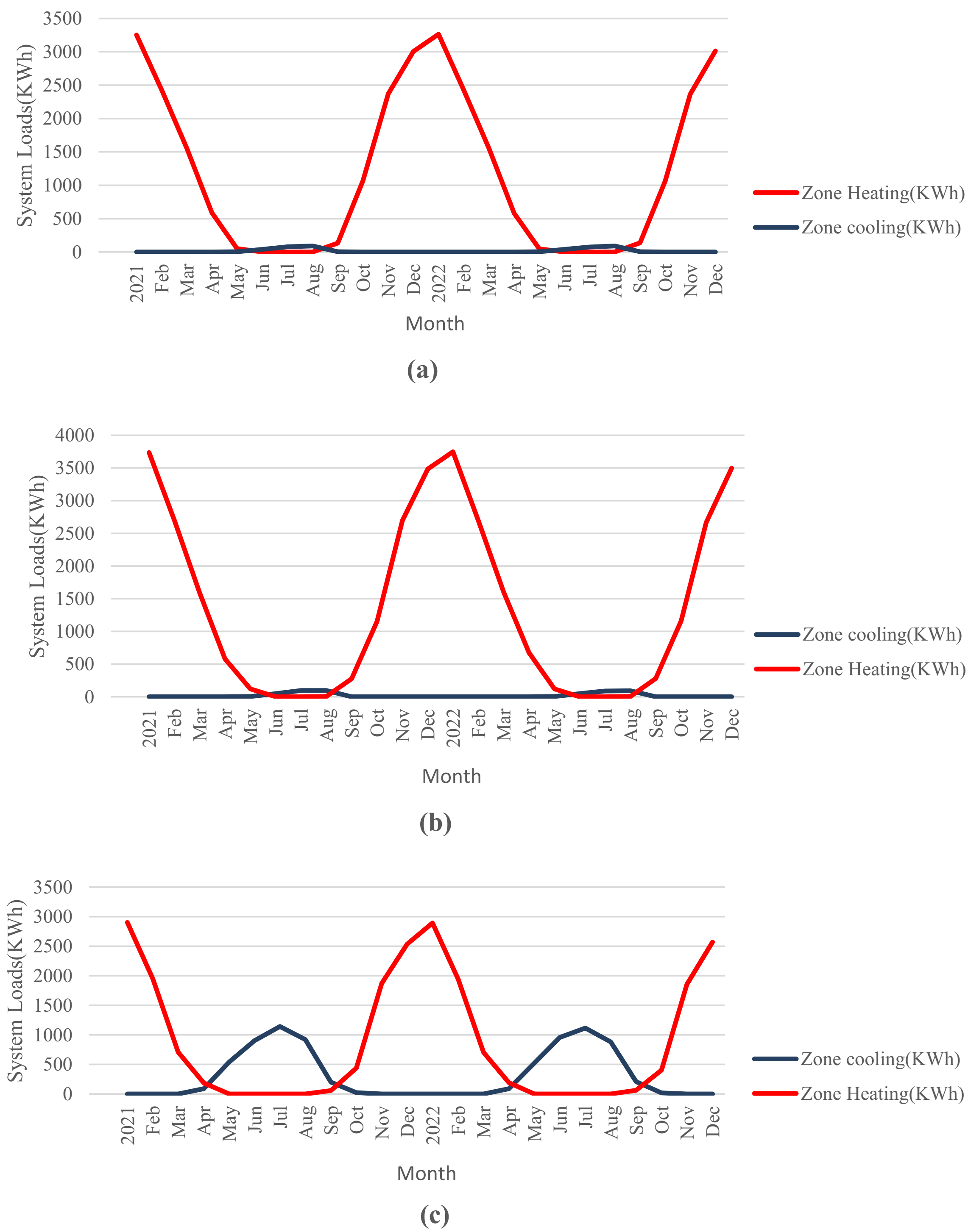


Fig. 12 **a** Cooling and heating load in a north–south building **b** Cooling and heating load in a south-north building **c** Cooling and heating load in a east–west building

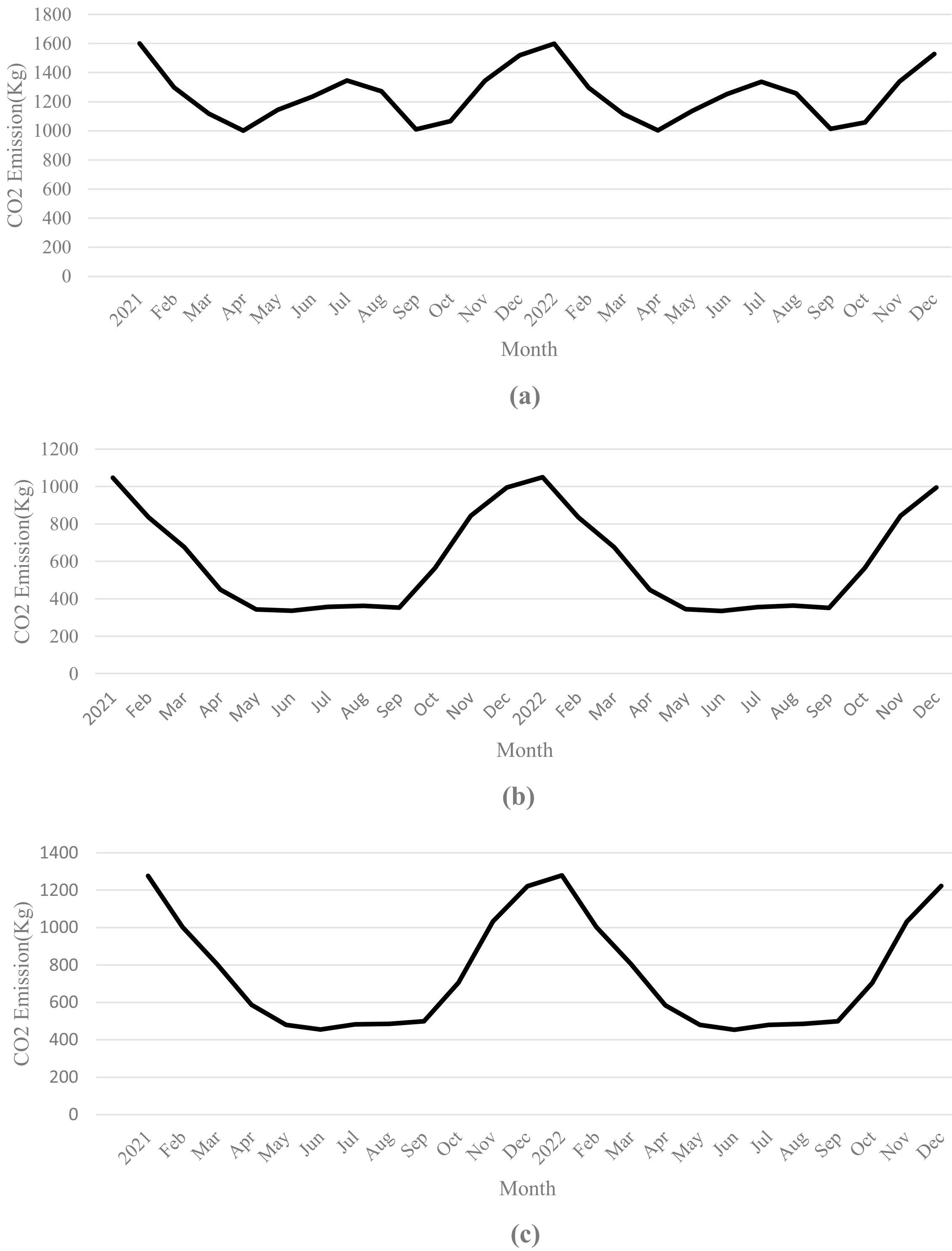


Fig. 13 **a** CO_2 emissions in a north–south building **b** CO_2 emissions in a south-north building (v) CO_2 emissions in a east–west building

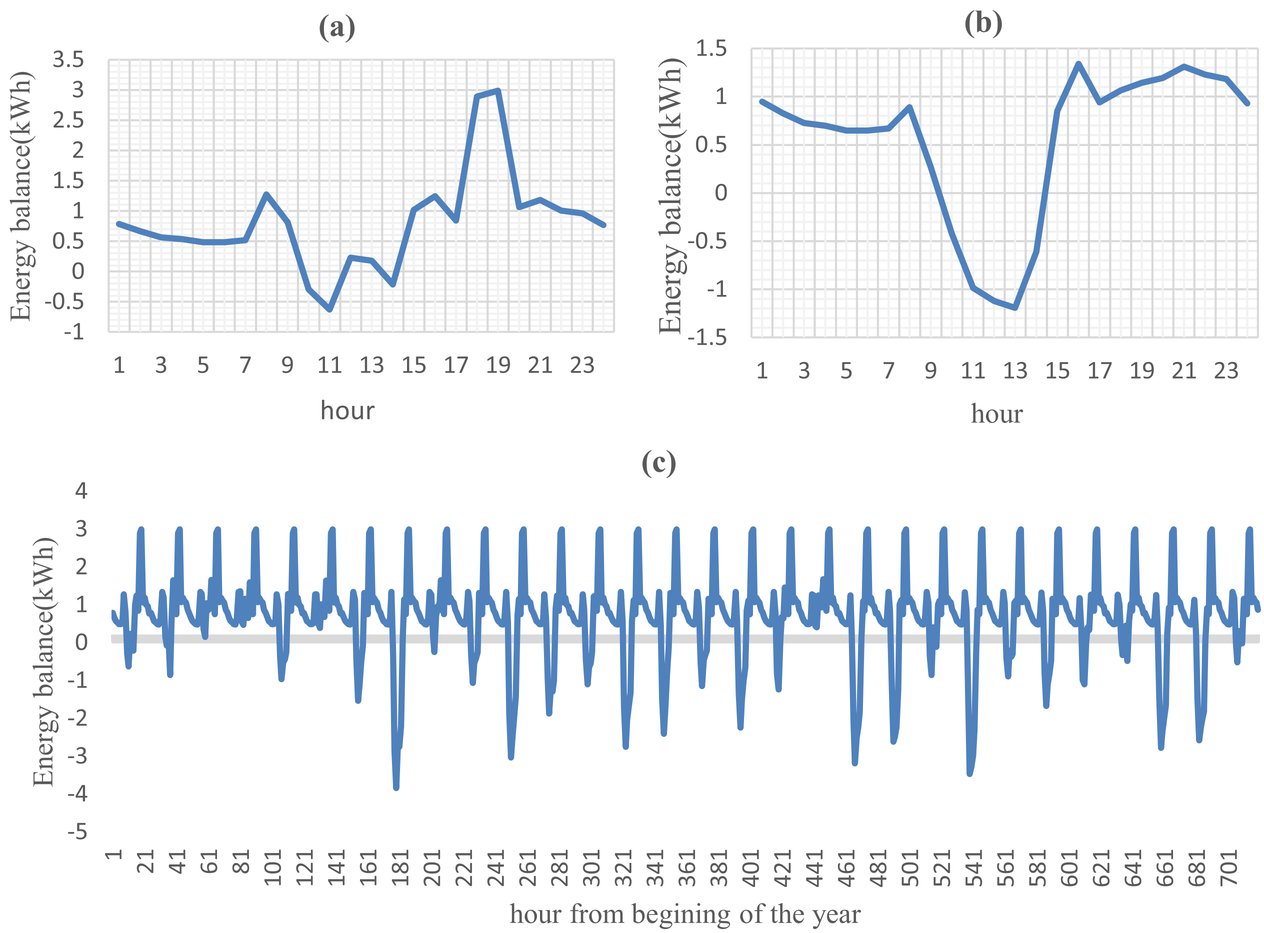


Fig. 14 **a** House Energy Balance No. 3904 on the 1 day of the year **b** House Energy Balance No. 3928 on the 1 day of the year **c** House Energy Balance No. 3904 for the 1 month of the year

Given that the building's area in the neighbourhood is almost the same and their maps are slightly different from each other, so it was predicted that the buildings in a row with the same geographical angle had approximately the same outputs and the main difference was when the geographical angles of the buildings were different.

Figure 12a shows that the need for cooling load in the Canadian climate is minimal, and the heating peak occurs in January, which is about 3000 kWh. As shown in Fig. 12b, north–south buildings, like south-north buildings, do not require cooling loads. The heating load in these buildings is slightly higher than in south-north buildings. Figure 12c shows that East–West buildings need cooling load due to their location in summer, and this peak occurs in July.

Figures 13a–c also show the production of CO2 for each of these categories. This produced carbon dioxide is proportional to the heating and cooling load of the building.

Energy balance results

To calculate the energy balance of each house, we subtract the hourly power generation data from the photovoltaic panels from the power consumption data from the simulation in the Builder Design software. The number is the energy balance of each house in a specific month, day, and hour of the year. It can be seen that there is a negative balance in some hours of the day. This means that electricity generation exceeds the consumption of the desired home.

Figure 14a and b show the energy balance for houses No. 3904 and 3928 on the first day of the year, January 1. It can be seen that the energy balance of each house varies according to its characteristics; Therefore, there is a need for a model that can predict the energy balance of the house according to the desired characteristics. Figure 14c shows the energy balance of house No. 3904 for the month of January, which varies with the hourly energy balance during 1 month of the year, considering

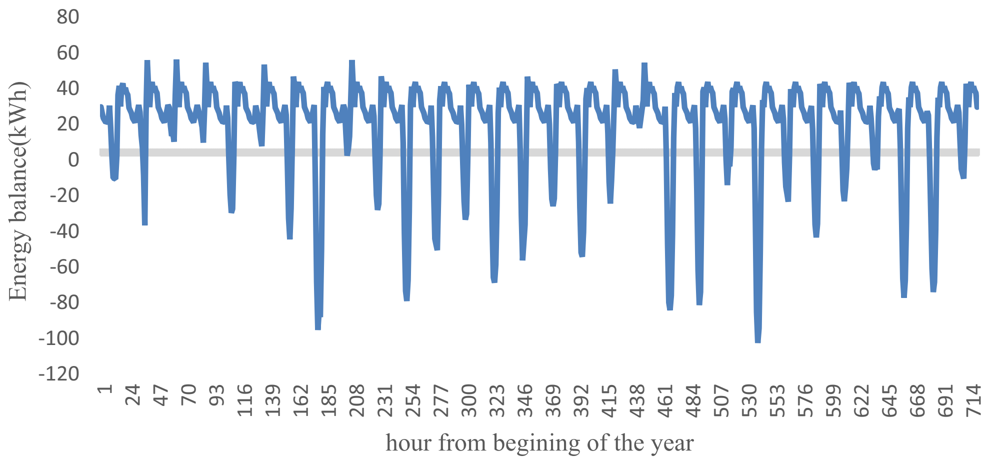


Fig. 15 Neighbourhood energy balance in the 1 month of the year

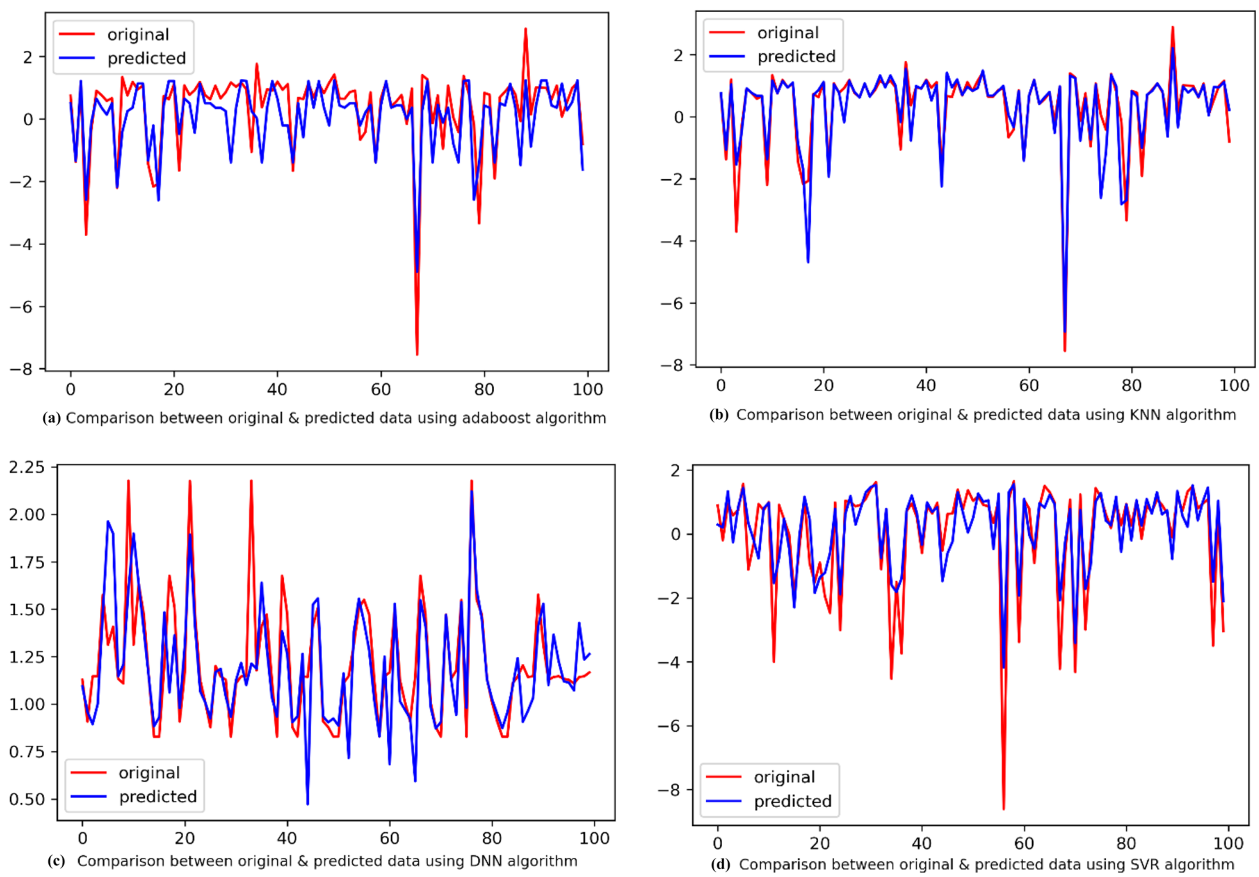


Fig. 16 Sample outputs of learning algorithms: **a** AdaBoost algorithm output **b** KNN algorithm output **c** DNN algorithm output **d** SVR algorithm output

the electricity consumption, intensity, and angle of the sun.

Figure 15 shows the energy balance of the study neighbourhood in January. According to the obtained data, it can be seen that the energy balance of the neighbourhood depends on the composition of the studied houses, so the focus of modeling should be on predicting the load of each house. Finally, after predicting the energy balance

Table 1 Algorithms RMSE

Algorithm	KNN	DNN	SVR	adaboost
MSE	0.56	0.53	0.92	0.95

of each house, the energy balance of the small network can be obtained.

Machine learning algorithm result

The developed data were given as input to the SVR, Adaboost, KNN, and DNN algorithms as seven input characteristics: hour, day, month, building type, solar panel surface area, building area of each building, and building heat load. The output and objective function of all models considered the energy balance of each house. For the KNN algorithm, the optimal k number for this algorithm was calculated to be 3. For the 3-layer DNN algorithm, the input layer with 128 neurons was selected according to 7 features ($2^7 = 128$), the hidden layer with 64 neurons, and the output layer with one neuron. In this section, learning outputs are plotted for each algorithm. Figure 15 shows the diagrams for all four algorithms.

In Fig. 16, the blue lines indicate the predicted results, and the red lines represent the actual outputs of the original data. For each algorithm, 100 test data were randomly selected from the database, and the actual value was compared with the value predicted by the corresponding model. As can be seen, the energy balance of each algorithm has a different accuracy. In the RMSE benchmark section, the accuracy of each algorithm is compared, and the best is recommended.

The resulting trained model can be developed in the neighbourhood and input a set of buildings with the features required for the model, and the output of the model will be the energy balance of all buildings during the day. The sum of all these balances reports the microgrid (neighbourhood) balance.

RMSE benchmark

Before entering the data into the machine learning algorithm, the data were divided into two groups, according to which 80% of the data set was used for training, and the other 20% were divided as testing data groups. Several models with different adjustment parameters were created during model training for each method. After regular adjustment to the maximum relevant parameters, each model was evaluated based on Root Mean Square Error (RMSE). Table 1 shows the RMSE value for each algorithm.

Conclusion

Due to the increasing emissions in cities, the need to replace electric vehicles and use solar panels as clean energy is increasingly felt. A model by which the energy balance of a home can be predicted by considering electric vehicles and solar panels is also of great importance. It can help study electric car tires and solar panels on the power grid. This paper extracted the data required for different machine learning algorithms by studying a neighbourhood in Edmonton, Canada, examining the different dimensions of houses in this neighbourhood, and simulating them.

This study analyzed a neighbourhood consisting of 29 houses in Edmonton, Canada, and a model related to the house electrical energy balance was developed. The most important parameters required for this modeling are the type of building, the area of the house infrastructure, the area of the house's photovoltaic panels, the house's heat load, and the time variables of the month, day, and hour. In this study, four well-known regression algorithms were used in the machine learning discussion, the best of which is the DNN algorithm with 128 neurons in the input layer, 64 neurons in the middle layer, and one neuron in the output layer. The predicted energy balance results show that during the hours of the day, when the sunlight is at its best in terms of radiation level and the radiation angle, the neighbourhood can generate more electricity than its consumption. Excess electricity can be delivered to the national grid. Due to the average production time of consumption, which is around 10 to 15 o'clock, this surplus electricity can be consumed in industries, which itself is a significant help in reducing the amount of carbon dioxide caused by electricity generation.

The algorithms for learning are SVR, Adaboost, KNN, and DNN models. Among these four algorithms, the DNN algorithm had the lowest RMSE index and using this model, a good prediction for neighbourhood balance was made. Further studies in this field are needed by considering more parameters and features of the building to consider a comprehensive model for green buildings and make this information available to decision-makers in this area to see an increasing number of homes using clean energy.

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by MAM and Mohammad Soleimani. The first draft of the manuscript was written by RZ and all authors commented on previous versions of the manuscript. AA supervised the manuscript. All authors read and approved the final manuscript.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Availability of data and materials
Datasets analyzed during the current study are available and can be given following a reasonable request from the corresponding author.

Declarations

Ethics approval and consent to participate
The authors herewith do confirm that this manuscript has not been published elsewhere and is not also under consideration by the other journals. The authors approve the presented manuscript and do agree with the submission under your management as the editor in chief of Journal of Sustainable Energy Research. The current study was carried out under the University of Tehran, Department of Energy Systems Engineering, Tehran, Iran.

Consent for publication
Not applicable.

Competing interests
The authors have no relevant financial or non-financial interests to disclose.

Author details
[1]Department of Renewable Energy and Environmental, University of Tehran, Tehran, Iran.

Received: 4 February 2023 Accepted: 15 May 2023
Published online: 19 June 2023

References
All Types of houses in Canada | WOWA.ca. Accessed: April. 06, 2023. https://wowa.ca/types-of-house-in-canada.
C. R. E. Association. (2019) National survey report of pv power applications in Canada.
Bui, D.-K., Nguyen, T. N., Ngo, T. D., & Nguyen-Xuan, H. (2020). An artificial neural network (ANN) expert system enhanced with the electromagnetism-based firefly algorithm (EFA) for predicting the energy consumption in buildings. *Energy, 190*, 116370.
Buonomano, A., Calise, F., Cappiello, F. L., Palombo, A., & Vicidomini, M. (2019). Dynamic analysis of the integration of electric vehicles in efficient buildings fed by renewables. *Applied Energy, 245*, 31–50.
City of Edmonton. (2017) Edmonton's green home guide.
Coffman, M., Bernstein, P., & Wee, S. (2017). Integrating electric vehicles and residential solar PV. *Transport Policy, 53*, 30–38.
J. Conti, P. Holtberg, J. Diefenderfer, A. LaRose, J. T. Turnure, and L. Westfall. (2016) International energy outlook 2016 with projections to 2040.
Deb, C., Zhang, F., Yang, J., Lee, S. E., & Shah, K. W. (2017). A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews, 74*, 902–924.
Dongare, A. D., Kharde, R. R., & Kachare, A. D. (2012). Introduction to artificial neural network. *Internal Journal Engineering Innovatons Technology, 2*(1), 189–194.
UN Env and IEA. (2017) Towards a zero-emission, efficient, and resilient buildings and construction sector.
Gholinejad, H. R., Adabi, J., & Marzband, M. (2022). Hierarchical energy management system for home-energy-hubs considering plug-in electric vehicles. *IEEE Transactions on Industry Applications, 58*(5), 5582–5592.
G. Göhler, F. Otteny, J. Triebke, & M. Reiser. (2019) Load profile generator for electric vehicle home charging," in *Lyon, France: 32ndElectric Vehicle Symposium (EVS32)*, pp. 19–22.
A. Gonçalves. (2018) Are electric cars (EC) really greener and eco-friendly. https://youmatter.world/en/are-electric-cars-eco-friendly-and-zero-emission-vehicles-26440/.
Guzel, I., & Göl, M. (2021). "Plug-in electric vehicle load modeling for smart charging strategies in microgrids", in. *International Conference on Smart Energy Systems and Technologies (SEST), 2021*, 1–6.
Hampshire, K., German, R., Pridmore, A., & Fons, J. (2018). Electric vehicles from life cycle and circular economy perspectives. *Version, 2*, 25.
"HousePlans. https://www.houseplans.pro/plans/plan/d-593.
Hou, X., Wang, J., Huang, T., Wang, T., & Wang, P. (2019). Smart home energy management optimization method considering energy storage and electric vehicle. *IEEE Access, 7*, 144010–144020.
Jain, R. K., Smith, K. M., Culligan, P. J., & Taylor, J. E. (2014). Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy. *Applied Energy, 123*, 168–178.
Khan, S., Sudhakar, K., & Yusof, M. H. (2022). Building integrated photovoltaics powered electric vehicle charging with energy storage for residential building: Design, simulation, and assessment. *Journal Energy Storage, 63*, 107050. https://doi.org/10.1016/j.est.2023.107050
Kim, T.-Y., & Cho, S.-B. (2019). Predicting residential energy consumption using CNN-LSTM neural networks. *Energy, 182*, 72–81. https://doi.org/10.1016/j.energy.2019.05.230
Lirola, J. M., Castaneda, E., Lauret, B., & Khayet, M. (2017). A review on experimental research using scale models for buildings: Application and methodologies. *Energy Building, 142*, 72–110.
Liu, T., Tan, Z., Xu, C., Chen, H., & Li, Z. (2020a). Study on deep reinforcement learning techniques for building energy consumption forecasting. *Energy Building, 208*, 109675.
Liu, Y., Chen, H., Zhang, L., Wu, X., & Wang, X. (2020). Energy consumption prediction and diagnosis of public buildings based on support vector machine learning: A case study in China. *J Cleaner Production, 272*, 122542. https://doi.org/10.1016/j.jclepro.2020.122542
Luo, X. J., Oyedele, L. O., Ajayi, A. O., & Akinade, O. O. (2020). Comparative study of machine learning-based multi-objective prediction framework for multiple building energy loads. *Sustain Cities Society, 61*, 102283. https://doi.org/10.1016/j.scs.2020.102283
Mokhtara, C., Negrou, B., Bouferrouk, A., Yao, Y., Settou, N., & Ramadan, M. (2020). Integrated supply–demand energy management for optimal design of off-grid hybrid renewable energy systems for residential electrification in arid climates". *Energy Conversion Management, 221*, 113192. https://doi.org/10.1016/j.enconman.2020.113192
Mokhtara, C., Negrou, B., Settou, N., Settou, B., & Samy, M. M. (2021). Design optimization of off-grid hybrid renewable energy systems considering the effects of building energy performance and climate change: case study of Algeria". *Energy, 219*, 119605. https://doi.org/10.1016/j.energy.2020.119605
Moss, R. H., et al. (2010). The next generation of scenarios for climate change research and assessment. *Nature, 463*(7282), 747–756.
Neves, S. A., Marques, A. C., Fuinhas, J. A., et al. (2017). Is energy consumption in the transport sector hampering both economic growth and the reduction of CO2 emissions? A disaggregated energy consumption analysis. *Transport Policy, 59*, 64–70.
Nutkiewicz, A., Yang, Z., & Jain, R. K. (2018). Data-driven Urban Energy Simulation (DUE-S): A framework for integrating engineering simulation and machine learning methods in a multi-scale urban energy modeling workflow. *Applied Energy, 225*, 1176–1189. https://doi.org/10.1016/j.apenergy.2018.05.023
O'Dwyer, E., Pan, I., Acha, S., & Shah, N. (2019). Smart energy systems for sustainable smart cities: Current developments, trends and future directions. *Applied Energy, 237*, 581–597.
Pfenninger, S., & Staffell, I. (2016). Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy, 114*, 1251–1265.
Quddus, M. A., Shahvari, O., Marufuzzaman, M., Usher, J. M., & Jaradat, R. (2018). A collaborative energy sharing optimization model among electric vehicle charging stations, commercial buildings, and power grid. *Applied Energy, 229*, 841–857.
Ramadhani, U. H., Shepero, M., Munkhammar, J., Widén, J., & Etherden, N. (2020). Review of probabilistic load flow approaches for power distribution systems with photovoltaic generation and electric vehicle charging". *International Jornal Electrical Power Energy System, 120*, 106003.
Savvides, A., Vassiliades, C., Michael, A., & Kalogirou, S. (2019). Siting and building-massing considerations for the urban integration of active solar energy systems. *Renewable Energy, 135*, 963–974.
Sawhney, A., & Kahn, M. E. (2012). Understanding cross-national trends in high-tech renewable power equipment exports to the United States. *Energy Policy, 46*, 308–318.

"Solar-Myheat." https://solar.myheat.ca/edmonton/.
Tian, W., et al. (2018). A review of uncertainty analysis in building energy assessment. *Renewable and Sustainable Energy Reviews, 93*, 285–301.
Wang, W., Hong, T., Xu, X., Chen, J., Liu, Z., & Xu, N. (2019). Forecasting district-scale energy dynamics through integrating building network and long short-term memory learning algorithm. *Applied Energy, 248*, 217–230.
Xiao, J., Li, Y., Xie, L., Liu, D., & Huang, J. (2018). A hybrid model based on selective ensemble for energy consumption forecasting in China. *Energy, 159*, 534–546.
Xu, X., Wang, W., Hong, T., & Chen, J. (2019). Incorporating machine learning with building network analysis to predict multi-building energy use. *Energy Buildings, 186*, 80–97. https://doi.org/10.1016/j.enbuild.2019.01.002
Zafar, B., & Ben Slama, S. A. (2022). PV-EV integrated home energy management using vehicle-to-home (V2H) technology and household occupant behaviors. *Energy Strategy Review, 44*, 101001. https://doi.org/10.1016/j.esr.2022.101001
Zhang, X., Chan, K. W., Li, H., Wang, H., Qiu, J., & Wang, G. (2020). Deep-learning-based probabilistic forecasting of electric vehicle charging load with a novel queuing model. *IEEE Trans. Cybern., 51*(6), 3157–3170.

Publisher's Note